Calculus, Algebra, and Analysis for JMC

Lectured by Marie-Amelie Lawn, Frank Berkshire

Typed by Aris Zhu Yi Qing

January 25, 2020

Contents

Chapter 1

Group theory

Study of the simplest algebraic structure on a set.

1.1 Basic Definitions and Examples

1.1.1 Binary operations and groups

Definition 1. *Set* is a collection of distinct elements. Let G be a set. ***Binary operation on G*** is a function

$$* : G \times G \to G(\text{Closure is included})$$

Example 2.

- $(\mathbb{N}, +), (\mathbb{Z}, +), (\mathbb{R}, \cdot)$
- $(\mathbb{N}, -)$ not a binary op. Not closed.
- $g, h \in G, g * h = h$
- Find a certain $c \in G$, define $g * h = c \; \forall g, h \in G$

Example 3. Cayley table: Draw a table of all the possible binary operations on a set. How many possible binary operations on a finite set with n elements? In general, there are ∞-many biniary operatiions. In this case, there are n^{n^2} possible binary operations. *In general,* $g_i * g_j \neq g_j * g_i$ (Not commutative!)

Definition 4. A binary operation $*$ on a set G is called associative if

$$(g * h) * k = g * (h * k) \; \forall g, h, k \in G$$

Example 5.

- $+$ on $\mathbb{N}, \mathbb{Z}, \mathbb{R}$? Yes
- $-$ on $\mathbb{R}$? No
- $g * h = g^h$ on $\mathbb{N}$? No

Definition 6. A binary operation is called commutative if

$$\forall g, h \in G, g * h = h * g$$

Example 7.

- $+, \cdot$ on $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$
- matrix multiplication ($AB \neq BA$ in general for A, B in $M(\mathbb{R}^n)$)
- let $g, h \in \mathbb{R}$, $g * h = 1 + g \cdot h$: commutative but *not associative*!

Definition 8. Let $(G, *)$ be a set. An element e is called *left identity* (respectively *right identity*) if:

$$e * g = g (\text{resp. } g * e = g) \; \forall \; g \in G$$

Caution: There might be *many* left/right identities or none.

Example 9.

1. let $(G, *)$ be a set with $g * h := g$. Find the left/right identities.

   ∞-many (or equal to the number of elements) right identities since h satisfies definition $\forall h$. No left identities: wanted $e * g = g = e$ by definitioin of $*$ (*unless only one element*).

2. $(G, *)$, $g * h = 1 + gh$. Ex: No right/left identities.

Idea: We want a good unique identity.

Theorem 10. let $(G, *)$ be set, such that $*$ has both a left identity e_1 *and* a right identity e_2, then

$$e_1 = e_2 =: e \quad \text{and} \quad e \text{ is unique.}$$

Proof.

- $e_1 = e_2$

$$\Rightarrow \left\{ \begin{array}{l} e_1 * g = g \Rightarrow e_1 * e_2 = e_2 \\ g * e_2 = g \Rightarrow e_1 * e_2 = e_1 \end{array} \right\} \forall\, g \in G \Rightarrow \; e_1 = e_2$$

- Unicity: Assume there exists another identity e'.

$$\Rightarrow e' * g = g * e' = g$$

$$e' * g = e' * e = e$$

$$g * e' = e * e' = e'$$

Therefore

$$e = e'$$

□

As soon as you get one left and one right identity, you have a unique identity e.

Definition 11. let $(G, *)$ be a set. Let $g \in G$. An element $h \in G$ is called left (resp. right) inverse if

$$h * g = e \text{ (resp. } g * h = e)$$

Caution: Again inverses might not exist, there might be many, or *not* the same on both sides.

Example 12.

(1) $(\mathbb{N}, \cdot)$ 1 has an inverse, otherwise *no* inverse.

(2) Find a binary operation on a set of 4 elements with left/right inverses not the same but identity e.

Theorem 13. Let $(G, *)$ be a set with associative binary operation and identity e. Then if h_1 is left inverse, and h_2 is right inverse, then

$$h_1 = h_2 = g^{-1} \text{ and } \textit{it is unique}$$

Proof.

- $h_1 = h_2$

  $h_1 * g = e, g * h_2 = e$. Therefore

  $$h_2 = e * h_2 = (h_1 * g) * h_2 = h_1 * (g * h_2) = e = h_1$$

- unicity: Assume $\exists g'^{-1}$ another inverse.

  $$g'^{-1} = e * g'^{-1} = (g^{-1} * g) * g'^{-1} = g^{-1} * (g * g'^{-1}) = g^{-1} * e = g^{-1}$$

□

(Group) Definition 14. A set $(G, *)$ with binary operatioin $*$ is called a *group* if:

(1) $*$ is associative

(2) $\exists e \in G$ an identity $\forall g \in G$

(3) All elements $g \in G$ have an inverse g^{-1}

Attention: The identity and inverses are *unique* by our previous results.

Example 15.

- $(\mathbb{Z}, +), (\mathbb{Z}_n, +)$ (will see this later) are groups.
- $(\mathbb{N}, +)$ not a group $\Rightarrow$ no inverses.
- $(\mathbb{C}, \cdot)$ not a group (0 has no multiplicative inverse), but $(\mathbb{C}^*, \cdot)$ is. ($\mathbb{C}^* = \mathbb{C} \backslash \{0\}$)
- $(G = \{e\}, *)$ with $e * e = e$ is a group called the *trivial group*.
- Empty set $\varnothing$ is not a group (No identity element.)

Definition 16. Let G be a group. It is called finite if it has finitely many elements.

Notation: $|G| = n$ (number of elements)

We say that G has ***order*** n. If $|G| = \infty$, the G is called an infinite group.

Example 17.

- the trivial group is finite, $|G| = 1$
- let $G = \{1, -1, i, -i\} \subset \mathbb{C}$, with $* = \cdot$. Is it a group? Yes. Check associativity, identity, and inverses.

(Abelian Group) Definition 18. A group is called *Abelian* if $*$ is commutative.

Example 19.

- previous example, trvial group, $(\mathbb{Z}, +), (\mathbb{C}^*, \cdot)$
- let $GL(\mathbb{R}^n)$ be the set of all invertible $n \times n$ matrices, $*$ = matrix multiplication. It is associative: $(AB)C = A(BC)$; It has identity: I_n. It has inverses: yes since we asked for it. So this is a group of matrices. But this is not Abelian since $AB \neq BA$.
- let G be the set of *invertible* functions with $* = \circ$, the composition of functions. Identity is $F(x) = x$; they are associative, invertible, but *not Abelian*.

1.1.2 Consequences of the axioms of group

Theorem 20. Let $(G, *)$ be a group, $g, h \in G$. Then

$$(g * h)^{-1} = h^{-1} * g^{-1}$$

Proof. To show: $(g * h) * (h^{-1} * g^{-1}) = e$.

Using assocativity, we have

$$g * (h * h^{-1}) * g^{-1} = g * g^{-1} = e$$

□

Definition 21. Let $n \in \mathbb{Z}$, let $(G, *)$ be a group and let $g \in G$. Then we definie g^n as follows:

$$g^n = \begin{cases} g * g * \cdots * g & n > 0 \\ g^{-1} * g^{-1} * \cdots * g^{-1} & n < 0 \\ e & n = 0 \end{cases}$$

where in the first case there are n copies of g in the product and ni the second there are $-n$ copies of g^{-1}, so that $g^n = (g^{-1})^{-n}$.

Theorem 22. Let $n, m \in \mathbb{Z}$ and let $G, *$ be a group. Then

1. $g^n * g^m = g^{n+m}$
2. $(g^n)^m = g^{nm}$

Proof. Exercise! (Hint: Induction.) □

1.1.3 Modular Artihmetic and the group $\mathbb{Z}_n$

Definition 23. let $n > 0$, $n \in \mathbb{Z}$ fixed, $a, b \in \mathbb{Z}$. a and b are called ***congruent modulo*** n if $n | a - b$.

Definition 24. $\forall a, b, c \in \mathbb{Z}$, $n > 0$ fixed in $\mathbb{Z}$:

(1) $a \equiv a \mod n$ (reflexivity)

(2) If $a \equiv b \mod n \iff b \equiv a \mod n$ (symmetry)

(3) if $a \equiv b \mod n$ and $b \equiv c \mod n \implies a \equiv c \mod n$ (transitivity)

Definition 25. Given a set S and an equivalence relation $\sim$ on S, the ***equivalence class*** of an element a in S is the set $\{x \in S \,|\, x \sim a\}$.

Definition 26. Define the equivalence class of $a \in \mathbb{Z}$ in the relation of congruence modulo n as:

$$[a]_n := \{b \in \mathbb{Z} \mid b \equiv a \mod n\}$$

Definition 27. Define equivalence classes $\mathbb{Z}_n$ as

$$\mathbb{Z}_n := \{[0]_n, [1]_n, \dots [n-1]_n\}$$

with 2 <u>binary operations</u> on $\mathbb{Z}_n$:

$+$: $\mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$, $([a]_n, [b]_n) \mapsto [a+b]_n$

$\cdot$: $\mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$, $([a]_n, [b]_n) \mapsto [ab]_n$

As we can see from the following lemma, the two operations are well-defined.

Lemma 28. Let $a, a', b, b' \in \mathbb{Z}$ s.t. $[a]_n = [a']_n, [b]_n = [b']_n$. Then $[a+b]_n = [a'+b']_n$, $[a \cdot b]_n = [a' \cdot b']_n$.

Proof. Exercise! □

Theorem 29. $(\mathbb{Z}_n, +)$ is an Abelian group.

Proof.

(1) Associativity:

$$\begin{aligned} ([a]_n + [b]_n) + [c]_n &= [a+b]_n + [c]_n \\ &= [a+b+c]_n \\ &= [a]_n + [b+c]_n \\ &= [a]_n + ([b]_n + [c]_n) \end{aligned}$$

(2) Commutativity:

$$\begin{aligned}[a]_n + [b]_n &= [a+b]_n \\ &= [b+a]_n \\ &= [b]_n + [a]_n\end{aligned}$$

(3) Identity element: $[0]_n$

(4) Inverse: Any element $[a]_n$ has an inverse $[-a]_n$.

$\square$

Example 30. $(\mathbb{Z}_n, \cdot)$ is an Abelian group?
Similary to above for associative, commutative, and identity.
Inverses:

Draw Caley table for $(\mathbb{Z}_3, \cdot)$. We realize that $[0]_3$ has no inverses. But $(\mathbb{Z}_3 \backslash \{[0]_3\}, \cdot)$ is.

Similarly, for $(\mathbb{Z}_4, \cdot)$, it does not have inverses for all classes.

Caution: In general $(\mathbb{Z}_n, \cdot)$ is *not* a group. The idea then is to make it a group by removing non-invertible elements.

Lemma 31. The element $[a]_n \in \mathbb{Z}_n$ has an inverse $\iff (a, n) = 1$.

Proof. $(a, n) = 1 \iff \exists b, c \in \mathbb{Z}$, s.t $ab + cn = 1 \iff cn = 1 - ab \iff \exists [b]_n$ s.t. $[a]_n[b]_n = [1]_n$. $\square$

Definition 32. $\mathbb{Z}_n^* := \{[a]_n \in \mathbb{Z}_n \mid \exists b \in \mathbb{Z} \text{ s.t. } [a]_n[b]_n = [1]_n\}$.

Theorem 33. $(\mathbb{Z}_n^*, \cdot)$ is an Abelian group.

Proof. To Show: if $[a]_n, [b]_n \in (\mathbb{Z}_n^*, \cdot) \Rightarrow [a]_n \cdot [b]_n \in (Z_n^*, \cdot)$.
$\Rightarrow (a, n) = (b, n) = 1 \Rightarrow (ab, n) = 1 \Rightarrow [ab]_n$ has inverse $[a]_n[b]_n$.

Alternatively: if g, h have inverse, $h^{-1}g^{-1}$ is inverse of gh. $\square$

1.2 Cyclic groups

Definition 34. Let G be a group, $g \in G$. The ***order*** of g is the smallest positive integer $n > 0$ such that $g^n = e$.

Notation: ord $g = n$. If $n = \infty$, then g is called of infinite order.

Example 35. $G = (\mathbb{C}^*, \cdot)$, ord $(-1) = 2$, ord $i = 4$, ord $2 = \infty$

Lemma 36. Let G be a finite group. Then every element $g \in G$ has finite orders.

Proof. Assume $g \in G$ has infinite orders. Write the list: $g^0, g^1, g^2, \ldots$

Since $|G| = n < \infty$, there are two elements g^k, g^l s.t. $g^k = g^l$, $k > l$. $\iff g^k g^{-l} = e \iff g^{k-l} = e$.

But then ord $g \leq k - l < \infty$. □

Lemma 37. Let G be a group, $g \in G$, ord $g = n$. Then all elements $\{g_0, g_1, g_2, \ldots, g^{n-1}\}$ are distinct.

Proof. Assume that $g^i = g^j$ for some $i, j, 0 \leq i \leq j \leq n-1$. Then $g^{j-i} = g^0 = e$. Since $i < j, j - i < n$. Since n is smallest integer, s.t. $g^n = e$, contradicts with the condition. □

Corollary 38. If $|G| = n < \infty$, $g \in G$, then ord $g \leq n$.

Proof. Assume $\exists i \in \mathbb{Z}, i \geq n+1,$ s.t. $g^i = e$ where $g \in G$, i is the smallest such integer. By previous lemma, $\{g_0, g_1, g_2, \ldots, g^{i-1}\}$ all distince. There are i elements $i > n$. □

Definition 39. We call a group G ***cyclic*** if

$$\exists g \in G \text{ s.t. } G = \{g^n | n \in \mathbb{Z}\}.$$

g is called a ***generator***.

Example 40.

- $(\mathbb{Z}, +)$. $2 = 1^2 = 1 + 1$, $n = 1^n$.
- $(\mathbb{Z}_n, +)$, generator $[1]_n$.
- $\{\pm 1, \pm i\}$, generator $\pm i$.

Lemma 41. All cyclic groups are Abelian.

Proof. To show: $\forall h, k \in G, h \cdot k = k \cdot h$.

G is cyclic $\Rightarrow G = \{g^n | n \in \mathbb{Z}\}$ for some generators $g \in G \Rightarrow h = g^i, k = g^j$.

$\Rightarrow h \cdot k = g^i \cdot g^j = g^{i+j} = g^{j+i} = g^j \cdot g^i = k \cdot h$. □

Warning: The converse *is not* true (Abelian does not imiply cyclic) One counter example is $(\mathbb{Q}, +)$. Assume $\mathbb{Q}$ is cyclic under $+$.

$$\Rightarrow \exists g \in \mathbb{Q} \text{ s.t. } q = g^n (= ng) \forall q \in \mathbb{Q}.$$

Take $\frac{g}{2}$ ($\in \mathbb{Q}$ since $g \in \mathbb{Q}$)

$$\Rightarrow \frac{g}{2} = ng \text{ for some } n \in \mathbb{Z}.$$

contradicting with original statements.

Lemma 42. Let G be a *finite* group, $|G| = n$. So

$$G \text{ is cyclic} \iff G \text{ contains an element of order } n$$

Proof.

"$\Rightarrow$": G is cyclic $\Rightarrow G$ has generator g. Assume ord $g = k$, so

$$\left\{g^0, \ldots, g^{k-1}\right\} \text{ are distinct.}$$

$\Rightarrow k = n$ since $|G| = n$.

"$\Leftarrow$": Let assume $\exists g \in G$, ord $g = n$.

$$\Rightarrow \left\{g^0, g^1, \ldots, g^{n-1}\right\} \text{ are all distinct.}$$

But $|G| = n$, hence g generates all the group. □

Lemma 43. Let G be a finite group. Then if G is cyclic, it has at most one element of order 2.

Proof. Since G is finite ($|G| = n$), and cyclic, $\exists g \in G$ of order n ($g^n = e$), and $G = \{g^0, g^1, \ldots, g^{n-1}\}$. Assume $\exists$ an element of order 2: $h = g^i, (i \geq 0, i \in \mathbb{Z})$, then

$$(g^i)^2 = e = g^{2i} \Rightarrow 2i = n \Rightarrow \begin{cases} n \text{ is even: exactly one element,} \\ n \text{ is odd: no element of order 2.} \end{cases}$$

□

Example 44. Are $(\mathbb{Z}_5^*, \cdot), (\mathbb{Z}_{15}^*, \cdot)$ cyclic? (Recall that the notation $\mathbb{Z}^* = \mathbb{Z}\backslash\{0\}$, and $\mathbb{Z}_n^*$ = set of all invertible congruence classes $[a]_n$.)

Hint: Use the previous lemma, or find out the generator.

1.3 Symmetric groups

1.3.1 Permutations

Definition 45. A function f from a set X to a set Y is called

- ***one-to-one*** or ***injective*** if $f(x_1) = f(x_2) \Rightarrow x_1 = x_2 \, \forall x_1, x_2 \in X$.
- ***onto*** or ***surjective*** if $\forall y \in Y, \exists x \in X$ s.t. $f(x) = y$.
- a ***bijection*** if it is both *injective* and *surjective*.

Furthermore, f is a bijection iif there is an inverse function $g : Y \mapsto X$ s.t. $g \circ f$ is the identity function on X and $f \circ g$ is the identity function on Y.

Definition 46. A ***permutation*** is a bijective function:

$$\sigma : \{1, 2, \ldots, n\} \mapsto \{1, 2, \ldots, n\}.$$

Notation: We write the permutation as *two-row notation*: we write down the numbers 1 to n, and underneath each number i we write down the number that σ sends i to:

$$\begin{vmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{vmatrix}$$

Because σ is a bijection, the bottom row of the table consists of the numbers $1, 2, \ldots, n$ in some order. So a permutation is a 're-ordering' of the numbers 1 to n.

Definition 47. The set of all permutation $S_n :=$ $\{\sigma : \{1, 2, \ldots, n\} \mapsto \{1, 2, \ldots, n\}\}$ is called the ***symmetric group*** (on n symbols).

Theorem 48. The set $(S_n, \circ)$ is a group.

Proof.

- Closure: Let $\nu, \tau \in S_n$, then ν, τ are bijective by definition, so are $\tau \circ \nu$ and $\nu \circ \tau$.
- Associativity: composition of functions is associative.
- Identity: identity $\nu(h) = k \; \forall k \in \{1, 2, \ldots, n\}$.
- Inverses: By definition: bijections $\iff$ $\exists$ inverses!

□

Theorem 49. $(S_n, \circ)$ is not Abelian.

Proof. Exercise! □

Proposition 50. $|S_n| = n!$.

Proof. Exercise! □

1.3.2 Cycle

Definition 51. A permutation is called a ***cycle*** if there is **a** sequence $\{a_1, a_2, \dots, a_k\}$ of distinct numbers s.t.

$$\sigma(a_1) = a_2, \quad \sigma(a_2) = a_3, \quad \dots, \quad \sigma(a_{k-1}) = a_k, \quad \sigma(a_k) = a_1$$

and $\sigma(i) = i$ for any other i *not* in the sequence. The number k is called the ***length*** of the cycle, and we often abbreviate 'cycle of length k' to 'k***-cycle***'.

Example 52.

$$\nu = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{vmatrix} \quad \text{and} \quad \tau = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{vmatrix}$$

ν is a 3-cycle, it rotates the numbers 1, 2, 3 and fixes 4. τ is not a cycle: no numbers are fixed, so if it was a cycle it would have to be 4-cycle, but it is not.

Proposition 53. The order of a k-cycle is k.

Proof. We know immediately that $\sigma^k = \mathrm{id}$ by definition. $\Rightarrow \operatorname{ord} \sigma \leq k$.

Assume that $\operatorname{ord} \sigma = i < k$. But by definition of $\sigma^i(a_1) = a_{i+1} \neq a_1$. □

Notation of a k-cycle: $(a_1, a_2, \dots, a_k)$. This means sending $a_1 \mapsto a_2 \mapsto a_3 \mapsto \cdots \mapsto a_k \mapsto a_1$ and fixes all other elements. This only makes sense if the numbers $a_1, a_2, \dots, a_k$ are all distinct (or this permutation would not be a cycle).

Example 54. From the previous example, we would write the 3-cycle ν as (1, 2, 3).

Note:

(1) There are several different ways of writing the same cycle, for instance (1, 2, 3), (2, 3, 1), (3, 1, 2) are all the same. The usual convention is to put the smallest number first.

(2) A cycle of length one has to be the identity permutation. So the 1-cycles (1), (3), (42), all denote the identity. The usual convention is to use (1), and this makes sense in any S_n.

(3) Cycles make sense if all elements are distince.

Example 55. The permutation $\tau \in S_4$ from the second previous example is not a cycle, but it is easy to see that it can be expressed as the composition

$$\tau = (3,4)(1,2)$$

of two 2-cycles.

Chapter 2

Applied Mathematical Methods

2.1 Differential Equations

2.1.1 Definitions and examples

Definition 56. An ***ordinary differential equation*** (ODE) for $y(x)$ is an equation involving derivatives of y.

$$f(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}, \frac{\mathrm{d}^2y}{\mathrm{d}x^2}, \ldots, \frac{\mathrm{d}^ny}{\mathrm{d}x^n}) = 0 \tag{2.1}$$

$$\frac{\mathrm{d}^ny}{\mathrm{d}x^n} = F(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}, \ldots, \frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}})$$

and we seek a solution (or solutions) for $y(x)$ satisfying the equations. (If there are more independent variables then we have a partial differential equation (PDE).)

Definition 57.

Order is the order of the highest derivative present.

Degree is the power of the highest derivative when fractional powers have been removed.

Linear *differential equation* is a differential equation that is defined by a *linear polynomial* in the unknown function and its derivative in each term of equation(2.1).

Example 58.

(a) Particle moving along a line with a given force $\rightarrow x(t)$ position as function of time t.

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = f\left(t, x, \frac{\mathrm{d}x}{\mathrm{d}t}\right)$$

e.g.

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\omega^2 x - 2k\frac{\mathrm{d}x}{\mathrm{d}t}$$

The first term is regarding the restoring force, while the second term is regarding the damping/friction. The function is of order 2, degree 1, and linear.

(b) Radius of curvature of a curve

It can be shown that

$$R(x, y) = \frac{\left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]^{\frac{3}{2}}}{\frac{\mathrm{d}^2y}{\mathrm{d}x^2}}$$

The function is of order 2 and degree 2.

(c) Simple growth and decay

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = kQ$$

The function is of order 1, degree 1, and linear. e.g.

(1) $k > 0$. Q as the quantity of money, and $k = (1 + \frac{r}{100})$, and r being the rate of interest.

(2) $k < 0$. Q as the amount of radioactive material, and k as the decay rate.

Hence, obviously $Q(t) = Q_0 e^{kt}$ where $Q_0 = Q(0)$ at $t = 0$.

(d) Population dynamics

$P(t)$ as population over time and $F(t)$ as food over time, with

$$\frac{\mathrm{d}P}{\mathrm{d}t} = aP(a > 0) \tag{2.2}$$

$$\frac{\mathrm{d}F}{\mathrm{d}t} = c(c > 0)$$

These two equations form a linear system, with both being of order 1, degree 1.

So $P(t) = P_0 e^{at}, F(t) = ct + F_0$. Misery! Population outgrows food supply.

Pierre Verhulst (1845) replaced a in equation(2.2) with $(a - bP)$ so that growth decreases as P increases:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = aP - bP^2 \tag{2.3}$$

This is in fact a *logistic ODE*, with order 1, degree 1, and nonlinear.

Note: Equation(2.3) is *separable*. Alternatively we can note that equation(2.3) is an example of a *Bernoulli differential equation*

$$\frac{\mathrm{d}y}{\mathrm{d}x} + F(x)y = H(x)y^n \tag{2.4}$$

with $n \neq 0, 1$ Substitution on $z(x) = (y(x))^{1-n} \Rightarrow$ a *linear* equation for $z(x) \to$ solution. (See below)

(e) Predator-Prey System

$x(t)$ as prey and $y(t)$ as predators, we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = ax - bxy, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = -cy + dxy \tag{2.5}$$

Note: Equation(2.5) is *separable* when written in principle

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}} \Rightarrow y(x) \Rightarrow x(t), y(t)$$

This is of order 1, degree 1, and a nonlinear system.

(f) Combat Model System

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -ay, \quad \frac{\mathrm{d}y}{\mathrm{d}t} = -bx \tag{2.6}$$

This is of order 1, degree 1, and linear system.

Note: Again equation(2.6) is *separable* when written as $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{bx}{ay} \Rightarrow y(x) \Rightarrow x(t), y(t)$

In general the solution of a differential equation of order n contains a number n of *arbitrary constants.* This general solution can be specialised to a particular solution by assigninig definite values to these constants.

Example 59.

(a) Family or parabolae $y = Cx^2$ as constant C takes different values.

On a particular curve of the family $\frac{\mathrm{d}y}{\mathrm{d}x} = 2Cx$. By substitutiion, eliminate $C \Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{2y}{x}$. This is a geometrical statement about slopes.

Note: 1st order differential equation $\leftrightarrow$ 1 arbitrary constant in general solution.

(b)

$$\left.\begin{aligned} x &= A\sin\omega t + B\cos\omega t \\ \frac{\mathrm{d}x}{\mathrm{d}t} &= A\omega\cos\omega t - B\omega\sin\omega t \\ \frac{\mathrm{d}^2x}{\mathrm{d}t^2} &= -A\omega^2\sin\omega t - B\omega^2\cos\omega t \end{aligned}\right\} \Rightarrow \frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \omega^2 x = 0$$

Note: 2nd order differential equation $\leftrightarrow$ 2 arbitrary constants in general solution.

Of course it's the reverse of this process we normally want to perform in order to get the general solution. We then often need a particular solution — which satisfieis certain other conditions — *boundary* or *initial condition.* These allow us to find the arbitrary constants in the solutions.

2.1.2 First Order Differential Equations

Properties and approaches

There are essentially 4 types we can solve *analytically*:

- ***separable***
- ***homogeneous***
- ***linear***
- ***exact*** (in Chapter "Partial Differentiation and Multivariable Calculus" later)

Let's look at them one by one:

(a) **Separable**

$$\frac{\mathrm{d}y}{\mathrm{d}x} = G(x) \cdot H(y)$$

Solve by rearrangement and integration

$$\int^y \frac{\mathrm{d}y}{H(y)} = \int^x G(x)\mathrm{d}x$$

E.g.

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= xy^2e^{-x} \\ \int \frac{1}{y^2}\mathrm{d}y &= \int xe^{-x}\mathrm{d}x \\ -\frac{1}{y} &= -xe^{-x} - e^{-x} + C \end{aligned}$$

Or singular solution $y = 0$.

If we want the particular solution which passes through $x = 1, y = 1$, then of course we need

$$C = -1 + 2e^{-1} \quad \text{and} \quad \frac{1}{y} = (x+1)e^{-x} + 1 - 2e^{-1}$$

(b) **Homogeneous**

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f\left(\frac{y}{x}\right)$$

Substitution $\frac{y}{x} = u(x)$, i.e. a new dependent variable,

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= u + x\frac{\mathrm{d}u}{\mathrm{d}x}(= f(u)) \quad \textit{\textbf{(Remember!)}} \\ f(u) - u &= \frac{x\mathrm{d}u}{\mathrm{d}x} \\ \int \frac{\mathrm{d}u}{f(u) - u} &= \int \frac{\mathrm{d}x}{x} \\ &\vdots \end{aligned}$$

E.g.

(i)

$$x^2\frac{\mathrm{d}y}{\mathrm{d}x} + xy - y^2 = 0$$
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{y}{x}\right)^2 - \frac{y}{x}$$
$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{u^2 - 2u}{x}$$
$$\vdots$$

(ii)

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x + y - 3}{x - y + 1}$$

This does not look homogeneous as it stands, but can be made so by substituting $x = 1 + X, y = 2 + Y$, and the expression becomes

$$\frac{\mathrm{d}Y}{\mathrm{d}X} = \frac{X + Y}{X - Y} = \frac{1 + \left(\frac{Y}{X}\right)}{1 - \left(\frac{Y}{X}\right)}$$

Then let $\frac{Y}{X} = u(X)$,

$$\Rightarrow \int \left(\frac{1 - u}{1 + u^2}\right) \mathrm{d}u = \int \frac{\mathrm{d}X}{X}$$

Eventually, the equation becomes

$$\tan^{-1}\frac{Y}{X} - \frac{1}{2}\ln\left(1 + \frac{Y^2}{X^2}\right) = \ln X + C$$
$$\tan^{-1}\left(\frac{y - 2}{x - 1}\right) - \frac{1}{2}\ln\left[(x - 1)^2 + (y - 2)^2\right] = C$$

Note: If we have e.g. $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x+y-3}{2(x+y)-7}$, then substitute $v(x) = x + y$ will work!

(c) **Linear**

$$\frac{\mathrm{d}y}{\mathrm{d}x} + F(x)y = G(x)$$

1st power only for y and $\frac{\mathrm{d}y}{\mathrm{d}x}$. We apply an *integrating factor* $R(x)$:

$$R(x) = \exp\left[\int^x F(x)\mathrm{d}x\right]$$

This allows us to form the expression

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[y\exp\left(\int^x F(x)\mathrm{d}x\right)\right] = G(x)\exp\left(\int^x F(x)\mathrm{d}x\right)$$

and then integrate...

E.g.

$$(x+2)\frac{\mathrm{d}y}{\mathrm{d}x} - 4y = (x+2)^6$$
$$\frac{\mathrm{d}y}{\mathrm{d}x} - \frac{4}{x+2} = (x+2)^5$$
$$\Rightarrow F(x) = -\frac{4}{x+2}, G(x) = (x+2)^5$$

Therefore,

$$R(x) = \exp\left[-\int^x \left(\frac{4}{x+2}\right)\mathrm{d}x\right] = \cdots = K(x+2)^{-4}$$

Subsequently, take $K = 1$ W.L.O.G.:

$$(x+2)^{-4}\frac{\mathrm{d}y}{\mathrm{d}x} - 4(x+2)^{-5}y = \frac{\mathrm{d}}{\mathrm{d}x}\left[y(x+2)^{-4}\right] = x+2$$

As such,

$$y(x+2)^{-4} = \frac{1}{2}x^2 + 2x + C \quad \text{(Put } C \text{ at the right time!)}$$
$$y(x) = \left(\frac{1}{2}x^{2+2x+C}\right)(x+2)^4$$

(So e.g. $y(0) = 8 \Rightarrow C = \frac{1}{2}$)

Novelties!

(i) Bernoulli equation (See Equation(2.4))
A nonlinear equation rendered linear by a substitution $u = y^{1-n}$...

(ii) E.g.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x+e^y}$$

It is <u>nonlinear</u> for $y(x)$ but <u>linear</u> for $x(y)$:

$$\frac{\mathrm{d}x}{\mathrm{d}y} - x = e^y \Rightarrow \ldots$$

2.1.3 'Special' Second Order Differential Equations

Definition 60. General Explicit form is

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = F\left(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

(a) $y, \frac{\mathrm{d}y}{\mathrm{d}x}$ **missing**, i.e.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = f(x)$$

Just integrate twice!

(b) $x, \frac{\mathrm{d}y}{\mathrm{d}x}$ **missing**, i.e.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = f(y)$$

Warning: Do not write $\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{1}{\frac{\mathrm{d}^2x}{\mathrm{d}y^2}}$. However, it may be true, but for what class of functions $y(x)$?

Let $\frac{\mathrm{d}y}{\mathrm{d}x} = p$,

$$\Rightarrow \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\mathrm{d}p}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = p\frac{\mathrm{d}p}{\mathrm{d}y} = \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{1}{2}p^2\right)$$

This substitution is effective because it eliminates x, so that the equation becomes separable for p and y.

Then we can integrate $\frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{1}{2}p^2\right) = f(y)$ w.r.t. y to get $p(y)$. Then using the definition of p,

$$x = \int \frac{\mathrm{d}y}{p(y)}$$

The same is obtained by multiplying the original equation by $\frac{\mathrm{d}y}{\mathrm{d}x}$ and recognizing $\frac{\mathrm{d}y}{\mathrm{d}x} \cdot \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{1}{2}\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]$

Example:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\omega^2 y$$

with ω being a real constant. (It is a simple harmonic motion.)

$$\Rightarrow \frac{1}{2}p^2 = -\frac{1}{2}\omega^2y^2 + C$$

Let $C = \frac{1}{2}\omega^2\overline{A}^2$. We therefore get

$$\frac{1}{p} = \frac{\mathrm{d}x}{\mathrm{d}y} = \pm\frac{1}{\omega(\overline{A}^2 - y^2)^{\frac{1}{2}}}$$

$$\begin{aligned}\Rightarrow \omega x + \overline{B} &= \pm \sin^{-1}\frac{y}{\overline{A}} \\ y &= \overline{A}\sin(\omega x + \overline{B}) \text{ W.L.O.G} \\ &= A\sin\omega x + B\cos\omega x\end{aligned}$$

(c) y **missing**, i.e.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = f\left(x, \frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

We put $\frac{\mathrm{d}y}{\mathrm{d}x} = p$, so

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = \frac{\mathrm{d}p}{\mathrm{d}x} = f(x,p)$$

i.e. First order $p(x)$. This substitution is effective because it eliminates y, so that the equation becomes <u>separable</u> for p and x.

Solve for $p(x)$ then integrate $\Rightarrow y(x)$.

<u>Example</u>: Radius of curvature

$$\frac{\left[1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right]^{\frac{3}{2}}}{\frac{\mathrm{d}^2y}{\mathrm{d}x^2}} = a \quad (a \text{ is an arbitrary constant})$$

$$\Rightarrow \frac{\mathrm{d}p}{\mathrm{d}x} = \frac{1}{a}(1+p^2)^{\frac{3}{2}}$$

$$\Rightarrow \frac{x}{a} + C = \int \frac{\mathrm{d}p}{(1+p^2)^{\frac{3}{2}}} \quad \text{i.e.} \quad \frac{x}{a} - \frac{A}{a} = \frac{p}{(1+p^2)^{\frac{1}{2}}}$$

$$\Rightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = p = \pm\frac{x-A}{\left[a^2 - (x-A)^2\right]^{\frac{1}{2}}}$$

$$\Rightarrow y = B \mp \left[a^2 - (x-A)^2\right]^{\frac{1}{2}} \quad \text{i.e.} \quad (x-A)^2 + (y-B)^2 = a^2$$

So they are all circles of radius a!

(d) x **missing**, i.e.

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = f\left(y, \frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

Yet again, let $\frac{\mathrm{d}y}{\mathrm{d}x} = p$, so

$$p\frac{\mathrm{d}p}{\mathrm{d}y} = f(y, p)$$

i.e. First order $p(y)$. So we solve for $p(y)$, then find $x = \int \frac{\mathrm{d}y}{p(y)}$.

Example:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = -\omega^2 y \mp 2k\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2$$

SHM with resistance proportional to $(\text{speed})^2$.
Hint: Solving this equation is the perfect application for solving Bernoulli Equation!

(e) **Linear Equations**, i.e. $y, \frac{\mathrm{d}y}{\mathrm{d}x}$ only occur to 1st power, if at all. So no products of y and $\frac{\mathrm{d}y}{\mathrm{d}x}$. The following section is dedicated to explaining the approach to solve linear differential equations.

General case — Linear Equations

The general form is, for order n,

$$\begin{aligned}\mathscr{L}y = a_0(x)\frac{\mathrm{d}^ny}{\mathrm{d}x^n} + a_1(x)\frac{\mathrm{d}^{n-1}y}{\mathrm{d}x^{n-1}} + a_2(x)\frac{\mathrm{d}^{n-2}y}{\mathrm{d}x^{n-2}} + \cdots \\ + a_{n-1}(x)\frac{\mathrm{d}y}{\mathrm{d}x} + a_n(x)y = f(x)\end{aligned} \tag{2.7}$$

where $a_0, a_1, \ldots, a_n$ and $f(x)$ are known functions of x only.
$\mathscr{L}$ is a **linear operator**, operating on $y(x)$:

$$\mathscr{L} \equiv \left[a_0\frac{\mathrm{d}^n}{\mathrm{d}x^n} + a_1\frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} + \cdots + a_n\right]$$

The equation(2.7) is called ***homogeneous*** iff $f(x) = 0$ and ***inhomogeneous*** iff $f(x) \neq 0$.
The homogeneous equation $\mathscr{L}y = 0$ has n independent solutions $y_1(x), y_2(x)$,

$\ldots, y_n(x)$ apart from *trivial* $y(x) = 0$. That is to say that $\mathscr{L}y_i(x) = 0$ for $i = 1, 2, \ldots, n$. (***Independence*** is an algebraic property. . .) Because of the linearity of $y_i(x)$ we find that the most general solution of the homogeneous equation $\mathscr{L}y = 0$ is given by

$$y(x) = A_1y_1(x) + A_2y_2(x) + \cdots + A_ny_n(x) \tag{2.8}$$

with $A_1, A_2, \ldots, A_n$ being arbitrary constants. This is because

$$\mathscr{L}y = \mathscr{L}\left(\sum_{i=1}^{n} A_iy_i(x)\right) = \sum_{i=1}^{n} A_i(\mathscr{L}y_i(x)) = 0$$

Of course equation(2.8) contains n arbitrary constants in accord with the order n of the differential equation.

For the inhomogeneous equation ($\mathscr{L}y = f(x)(2.7)$), the expression(2.8) is called the ***complementary functions*** (CF) of equation(2.7). Any solution of the inhomogeneous equation(2.7), say $Y(x)$, is called a ***particular integral*** (PI) of equation(2.7). The most general solution of equation(2.7) is thus

$$y(x) = (\mathrm{CF}) + (\mathrm{PI})$$

This contains n arbitrary constants as required/expected!

The constants can be specified in practice to produce a particular solution which satisfies (n) initial/boundary conditions.
<u>Note</u>

(a) For any two solutions $Y_1(x), Y_2(x)$ of equation(2.7), their difference satisfies

$$\mathscr{L}(Y_1 - Y_2) = \mathscr{L}Y_1 - \mathscr{L}Y_2 = f(x) - f(x) = 0$$

(b) Generally, finding $y_1(x), y_2(x), \ldots, y_n(x)$ functions might be very tough — our differential equation has generally variable coefficients after all! So we look at the most common case we need to study — constant coefficients! W.L.O.G.:

$$a_0(x) = 1, a_1(x) = a_1, a_2(x) = a_2, \ldots, a_n(x) = a_n$$

Linear Equations — Second Order, Constant Coefficients

Consider

$$\mathscr{L}y = \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_2 y = f(x) \tag{2.9}$$

Alternatively, in terms of notation,

$$\mathscr{L}y = y'' + a_1 y' + a_2 y = f(x)$$

Overall flow of solving the equation is to firstly find CF then PI,

$$\Rightarrow y(x) = \mathrm{CF} + \mathrm{PI}$$

Finding the CF We need to solve

$$\mathscr{L}y = \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_2 y = 0 \tag{2.10}$$

Try a solution of the form $y = e^{\lambda x}$ where λ is a constant — which we need to find! (It works by demonstration.) Evidently,

$$(\lambda^2 + a_1\lambda + a_2)e^{\lambda x} = 0$$

The exponential cannot help — for any λ let alone for all x. So

$$\lambda^2 + a_1\lambda + a_2 = 0 \tag{2.11}$$

as the auxiliary equations. In general, there are two distinct roots λ_1, λ_2 of this quadratic, so that $e^{\lambda_1 x}, e^{\lambda_2 x}$ are solutions of equation(2.10), i.e.

$$\mathscr{L}\left(e^{\lambda_1 x}\right) = 0 = \mathscr{L}\left(e^{\lambda_2 x}\right)$$

Because of the linearity property of $\mathscr{L}$ we have

$$y_{\mathrm{CF}} = A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x}$$

where A_1, A_2 are two arbitrary constants and $\mathscr{L}y_{\mathrm{CF}} = 0$ as required.

If the roots of (2.11) are equal, i.e. $\lambda_1 = \lambda_2 = \lambda$, then certainly $A_1 e^{\lambda x}$ is a solution of (2.10) with *one* arbitrary constant — we need *another*! A second linearly independent solution is given by $A_2 x e^{\lambda x}$, so that we have

$$y_{\mathrm{CF}} = A_1 e^{\lambda x} + A_2 x e^{\lambda x}$$

We can see this easily: (2.11) must take the form $(\lambda + \frac{a_1}{2})^2 = 0$ since $a_2 = \frac{a_1^2}{4}$ and $\lambda = -\frac{a_1}{2}$ (repeated root). Then substituting $xe^{\lambda x}$ into (2.10) we have

$$\mathscr{L}\left(xe^{\lambda x}\right) = (2\lambda + a_1)e^{\lambda x} + (\lambda^2 + a_1\lambda + a_2)xe^{\lambda x} = 0$$

as required. Here, n in $\mathscr{L}$ is 2.

Example 61.

1.
$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 5\frac{\mathrm{d}y}{\mathrm{d}x} + 6y = 0$$
$\Rightarrow \lambda^2 + 5\lambda + 6 = 0,\ \lambda = -3, -2.$ So
$$y(x) = A_1 e^{-3x} + A_2 e^{-2x}$$

2.
$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 4\frac{\mathrm{d}y}{\mathrm{d}x} + 4y = 0$$
$\Rightarrow \lambda^2 + 4\lambda + 4 = 0,\ \lambda = -2, -2.$ So
$$y(x) = A_1 e^{-2x} + A_2 x e^{-2x}$$

What about *complex roots* of (2.11)? (assuming $a_1, a_2 \in \mathbb{R}$) We know that the roots are complex conjugates, i.e. $\lambda_{1,2} = \alpha \pm i\beta, \alpha, \beta \in \mathbb{R}$. Now, formally our solution is, as above,

$$y = A_1 e^{(\alpha + i\beta)x} + A_2 e^{(\alpha - i\beta)x}$$

Since $\beta \neq 0$ here since the roots cannot be equal! so we can rewrite in alternative forms:

$$y = e^{\alpha x}\left[A_1 e^{i\beta x} + A_2 e^{-i\beta x}\right] = e^{\alpha x}\left[C_1 \cos \beta x + C_2 \sin \beta x\right]$$

where A_1, A_2 or C_1, C_2 can be taken as our arbitrary constants. (Naturally, $C_1 = A_1 + A_2, C_2 = (A_1 - A_2)i$ by De Moivre.)

Example 62.

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + 2k\frac{\mathrm{d}x}{\mathrm{d}t} + \omega^2 x = 0$$

which is the equation for damped harmonic oscillator ($k > 0$).

$$\lambda^2 + 2k\lambda + \omega^2 = 0, \quad \lambda_{1,2} = -k \pm \sqrt{k^2 - \omega^2}$$

and

$$x(t) = A_1e^{\lambda_1 t} + A_2e^{\lambda_2 t}$$

in general. This can be broken down into different cases.

(1) $k = 0$, i.e. *No Damping.*

$$x = A_1e^{i\omega t} + A_2e^{-i\omega t} = C_1\cos\omega t + C_2\sin\omega t$$

(2) $k^2 < \omega^2$, i.e. *Light Damping.*

$$x = A_1e^{-kt+i\omega t} + A_2e^{-kt-i\omega t} = (C_1\cos\omega t + C_2\sin\omega t)e^{-kt}$$

with $\omega = (\omega^2 + k^2)^{\frac{1}{2}}$.

(3) $k^2 > \omega^2$, i.e. *Heavy Damping.*

$$x = A_1e^{-|\lambda_1|t+A_2e^{-|\lambda_2|t}}$$

since λ_1, λ_2 are each neagative real.

(4) $k^2 = \omega^2$, i.e. *Critical Damping.*

$$\lambda_1 = \lambda_2 = -k \Rightarrow x = (A_1 + A_2t)e^{-kt}$$

Note: $x(t)$ behaviours for various cases!

Finding a PI Now we have the CF we need any particular solution of (2.9), in order to complete the job of finding the general solution. The PI is *not* unique! Our guide is the form of the function $f(x)$ on RHS.

(a) *polynomial in x*
Try a polynomial for the PI and choose the coefficients to fit! Example:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 3\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = x$$

Try PI $= ax^2 + bx + c$, where we need to find a, b, c. This method is often known as the method of undetermined coefficients.
We now determine them! (SIAS — Suck It And See)

$$2a - 3(2ax + b) + 2(ax^2 + bx + c) = x$$

By comparing the coefficients, we can obtain

$$a = 0, b = \frac{1}{2}, c = \frac{3}{4} \Rightarrow y_{\mathrm{PI}} = \frac{1}{2}x + \frac{3}{4}$$

Since $y_{\mathrm{CF}} = A_1e^x + A_2e^{2x}$ for this equation, then the general solution can be written as

$$y(x) = A_1e^x + A_2e^{2x} + \frac{1}{2}x + \frac{3}{4}$$

Note: Our inclusion of ax^2 term in our trial PI has been self-correcting since it emerged that $a = 0$. This is always so; the method gives what is needed!

(b) *multiple of e^{bx}*
The obvious choice for the PI is Ae^{bx}, since the linear operator $\mathscr{L}$ generates only terms of this type — choose A to fit! But there are two cases to consider:

(i) e^{bx} *not* in y_{CF}, i.e. $\mathscr{L}(e^{bx}) \neq 0$
Example:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 5\frac{\mathrm{d}y}{\mathrm{d}x} + 6y = 7e^{8x}$$

with

$$y_{\mathrm{CF}} = A_1e^{-3x} + A_2e^{-2x}$$

Try $y_{\mathrm{PI}} = Ae^{8x}$, then

$$Ae^{8x}[64 + 40 + 6] = 7e^{8x} \Rightarrow A = \frac{7}{110}$$

and general solution is

$$y(x) = y_{\text{CF}} + \frac{7}{110}e^{8x}$$

(ii) e^{bx} is *contained* in y_{CF}, i.e. $\mathscr{L}e^{bx} = 0$

Our trial solution in (i) now does not work! We might hope (anticipate) that xe^{bx} might be involved, and just try it… (SIAS)

A more 'automatic' approach is to take the Ae^{bx} from the CF (where A was constant) and try a PI of the form $A(x)e^{bx}$ — called ***variation of parameters***. We expect that $A(x)$ will be a polynomial in x!

Example:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + 3x + 2y = e^{-x}$$

with

$$y_{\text{CF}} = A_1e^{-x} + A_2e^{-2x}$$

Try $y_{PI} = A(x)e^{-x}$.

$$\Rightarrow (A'' - 2A' + A)e^{-x} + 3(A' - A)e^{-x} + 2Ae^{-x} = e^{-x}$$

By comparing the coefficients, we get

$$A'' + A' = 1$$

Afterwards, integrate with respect to x once and we get

$$A' + A = x + \overline{C_1}$$

Solving this first-order linear equation, and we get

$$A = x + C_1 + C_2e^{-x}$$

$$\Rightarrow y_{\text{PI}} = A(x)e^{-x} = xe^{-x} + C_1e^{-x} + C_2e^{-2x}$$

Take PI $= xe^{-x}$ (W.L.O.G), we can obtain

$$y(x) = A_1e^{-x} + A_2e^{-2x} + xe^{-x}$$

Of course if the auxiliary equation has equal roots then y_{CF} has xe^{bx} too! However the variabtion of parameters still works — or alternatively (a trial polynomial)(e^{bx}).

<u>Example</u>:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 4\frac{\mathrm{d}y}{\mathrm{d}x} + 4y = e^{-2x}$$

with

$$y_{\mathrm{CF}} = A_1 e^{-2x} + A_2 x e^{-2x}$$

We can then set PI as

$$y_{\mathrm{PI}} = A(x)e^{-2x} \Rightarrow \ldots A'' = 1 \Rightarrow A = \frac{x^2}{2} + [\overline{A_1} + \overline{A_2}x]$$

$$\Rightarrow y(x) = A_1 e^{-2x} + A_2 x e^{-2x} + \frac{x^2}{2}e^{-2x}$$

(c) e^{bx} is *polynomial* in x

Try PI $= C(x)e^{bx}$ where $C(x)$ is a polynomial with coefficients to be found — as in (a), (b) above.

(d) sines, cosines, sinh, cosh

We *either* just recognize the pattern and put e.g. $A\cos() + B\sin()$ *or* $A\cosh() + B\sinh()$, etc.

<u>*OR*</u>

Make use of exponentials — maybe complex ones using $e^{ix} = \cos x + i\sin x$, etc.

<u>Example</u>:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 3\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = e^x \cos x$$

with

$$y_{\mathrm{CF}} = A_1 e^{-x} + A_2 e^{-2x}.$$

There is no obvious trouble with this CF...

(1) Try $y_{\mathrm{PI}} = Be^x\cos x + Ce^x \sin x$ because $\mathscr{L}(y_{\mathrm{PI}})$ produces terms of a similar type. Substitute in and equate coefficents of $e^x\cos x, e^x \sin x$ on the two sides $\Rightarrow B = \frac{1}{10}, C = \frac{1}{10}$.

<u>*OR*</u>

(2) Put RHS $= \frac{1}{2}e^{(1+i)x} + \frac{1}{2}e^{(1-i)x} (= \Re\left(e^{(1+i)x}\right))$. Then try

$$y_{\mathrm{PI}} = C_1 e^{(1+i)x} \Rightarrow \left[(1+i)^2 + 3(1+i) + 2\right] C_1 = 1$$

and $C_1 = \frac{1}{5(1+i)} = \frac{1}{10}(1-i)$, and

$$y_{\mathrm{PI}} = \Re\left[\frac{1}{10}(1-i)e^{(1+i)x}\right] = \frac{1}{10}e^x \cos x + \frac{1}{10}e^x \sin x$$

Naturally, we might need to be adaptable if we find polynomials on RHS in $f(x)$ as well, or the 'equal roots' case...However something to beware:

<u>Example</u>:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 3\frac{\mathrm{d}y}{\mathrm{d}x} + 2y = \cosh 2x$$

with

$$y_{\mathrm{CF}} = A_1 e^{-x} + A_2 e^{-2x}$$

If we try $y_{\mathrm{PI}} = C_1 \cosh 2x + C_2 \sinh 2x$, we would find C_1, C_2 not defined...

$$\begin{cases} 6C_1 + 6C_2 = 1 \\ 6C_1 + 6C_2 = 0. \end{cases}$$

Why?! Well $\cosh 2x = \frac{1}{2}(e^{2x} + e^{-2x})$ and one of these exponentials *is* in y_{CF}. The better one is

$$y_{\mathrm{PI}} = \frac{1}{24}e^{2x} - \frac{1}{2}xe^{-2x}$$

using earlier results.

<u>Conclusion</u>: Try to use complex numbers, because it avoids "clashing" with hyperbolic functions, and also prevents calculation mistakes, like what would happen when differentiating sines and cosines.

Of course we might finally need to specialise our general solution to the particular solution that satisfies particular boundary conditions.

<u>Example</u>:

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + \frac{\mathrm{d}y}{\mathrm{d}x} - 6y = \sin x + xe^{2x}$$

subject to $y(0) = 0$, $\frac{\mathrm{d}y}{\mathrm{d}x}(0) = 0$. The general solution is

$$y(x) = A_1 e^{-3x} + A_2 e^{2x} - \frac{1}{50}(\cos x + 7\sin x) + \frac{e^{2x}}{50}(5x^2 - 2x)$$

and then

$$\left.\begin{aligned} 0 &= A_1 + A_2 - \frac{1}{50} \\ 0 &= -3A_1 + 2A_2 - \frac{7}{50} - \frac{1}{25} \end{aligned}\right\} \Rightarrow \begin{cases} A_1 = -\dfrac{7}{250} \\ A_2 = \dfrac{12}{250}. \end{cases}$$

2.1.4 Equations with variable coefficients

Special types to meet later (Bessel, Legendre, etc.) ...
A Novelty due to Euler (+ Cauchy!) If W.L.O.G.

$$x^n \frac{\mathrm{d}^n y}{\mathrm{d}x^n} + b_1 x^{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + b_n y = f(x)$$

with $b_1, b_2, \ldots, b_n$ constants.

(i) $f(x) = 0$. Try $y = x^\lambda \Rightarrow$ n values of λ in general.

$$y(x) = A_1 x^{\lambda_1} + A_2 x^{\lambda_2} + \cdots + A_n x^{\lambda_n}$$

with n arbitrary constants.

(ii) $f(x) \neq 0$. The method in (i) above might not be nice for PI! So put $x = e^t$ to *stretch* the independent variable, becoming a *linear equation for* $y(t)$ which has constant coefficients.

Example:

$$x^2 \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + 3x \frac{\mathrm{d}y}{\mathrm{d}x} + y = x^3.$$

Let $x = e^t$, so $\frac{\mathrm{d}x}{\mathrm{d}t} = e^t = t$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\frac{\mathrm{d}y}{\mathrm{d}t}}{\frac{\mathrm{d}x}{\mathrm{d}t}} = \frac{1}{e^t}\frac{\mathrm{d}y}{\mathrm{d}t}$$

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}y}{\mathrm{d}x}}{\frac{\mathrm{d}x}{\mathrm{d}t}} = \frac{\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-t}\frac{\mathrm{d}y}{\mathrm{d}t}\right)}{e^t} = -e^{-2t}\frac{\mathrm{d}y}{\mathrm{d}t} + e^{-2t}\frac{\mathrm{d}^2 y}{\mathrm{d}t^2}.$$

The equation therefore becomes

$$\left(\frac{\mathrm{d}^2y}{\mathrm{d}t^2} - \frac{\mathrm{d}y}{\mathrm{d}t}\right) + 3\frac{\mathrm{d}y}{\mathrm{d}t} + y = e^{3t}$$

i.e.

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + 2\frac{\mathrm{d}y}{\mathrm{d}t} + y = e^{3t}.$$

So

$$y(t) = A_1e^{-t} + A_2te^{-t} + \frac{1}{16}e^{3t}$$

and

$$y(x) = \frac{A_1}{x} + \frac{A_2}{x}\ln x + \frac{1}{16}x^3.$$

We should note that $x > 0$ and $x < 0$ need to be treated separately since $x = 0$ is an evident singularity. For $x < 0$ we would need to substitute $x = -e^t$ in the above method.

2.2 Difference Equations

2.2.1 Definitions and Examples

(Recurrence relations, maps, discrete dynamical systems, ...) From variables whose change is '*continuous*', we now consider variables which are '*discrete*'. ('Season to season', 'one accounrting period to the next', etc.) We have a *dependent variable* $U(n)$ with *integer independent variable* n — together with a relation connecting $U(n)$ to $U(n+1), U(n+2), \cdots$.

Note:

(i) ***Order*** corresponds to how many succeeding generations are involved.

(ii) ***Difference equation*** is associated with e.g. $A(n+1) - A(n) = f[A(n)]$, for instance.

Example 63.

(a) Fibonacci Sequence

Leonardo of Pisa wondered about how many rabbit pairs would be produced in the nth generation starting from a single pair and supposing that any pair from one generation produces a new pair each generation after an initial gap. . .

$$\begin{cases} U(n) & = 1 \quad 1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \cdots \\ n & = 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad \ \ 7 \cdots \end{cases}$$

and

$$U(n+2) = U(n) + U(n+1).$$

The equation is homogeneous (because only function of $U(n)$ is present without a single term of $f(n)$), linear, and second order.

(b) Money!

If we have an amount $A(n)$ at the beginning of an accounting period, then the amount at the end of that period (i.e. at the beginning of the next) is

$$A(n+1) = (1 + \frac{R}{100})A(n)$$

where $R\%$ is interest rate. The equation is homogeneous, linear, and first order.

If a payment is made each period, then

$$A(n+1) = (1 + \frac{R}{100})A(n) - P.$$

The equation is inhomogeneous, linear, and first order.

(c) Population Dynamics

Population $P(n)$ of an organism measured in each season is

$$P(n+1) = aP(n) - b[P(n)]^2$$

where a, b are positive. The first term indicates the growth, while the second term indicates the overcrowding or competition. (It is quadratic because it relates to the *interactions* of two entities, and the number of way to choose such as pair from a population is quadratic!)

This is a form of what is known as the ***logistic map***. It is homogeneous, nonlinear, anda first order. This turns out to have many different behaviours to that of logistic differential equation.

2.2.2 Linear Difference Equations

Broadly we use methods very similar to those we employed for linear differential equations — particularly terminologies like 'Complementary Function' and 'Particular integral', 'number of arbitrary constants', 'order', ...

Example 64.

(a) Fibonacci Sequence

$$U(n+2) = U(n) + U(n+1)$$

Try $U(n) = A\lambda^n$, where A is an arbitrary constant and λ is a particular constant (to be found). We can therefore obtain the *characteristic equation* (as compared with the *auxiliary equation* in differential equations):

$$\lambda^2 - \lambda - 1 = 0.$$

$$\Rightarrow \lambda_{1,2} = \frac{1}{2} \pm \frac{1}{2}\sqrt{5} = \tau, -\frac{1}{\tau}$$

with $\tau = 1.6180\ldots$, which is the golden number. We therefore get

$$\begin{aligned} U(n) &= A_1\lambda_1^n + A_2\lambda_2^n \\ &= A_1\tau^n + A_2\left(-\frac{1}{\tau}\right)^n. \end{aligned}$$

Substitute in $U(1) = 1, U(2) = 1$, we obtain $A_1 = \frac{1}{\sqrt{5}}, A_2 = -\frac{1}{\sqrt{5}}$.

$$\Rightarrow U(n) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

which is known as the "Binet formula".

A particular interesting identity as an application of the Fibonacci Sequence is the "Cassini's identity":

$$U(n+2)U(n) - [U(n+1)]^2 = (-1)^{n+1}$$

which can show that $13 \times 5 - 8^2 = 1$.

There are other sequences, such as the Lucas sequence, where $U(1) = 1, U(2) = 3$, etc.

(b) MoneyA

$$A(n+1) - \left(1 + \frac{R}{100}\right) A(n) = -P$$

$A(n)_{\text{CF}}$ is obtained by solving LHS $= 0$. Try

$$A(n) = A\lambda^n \Rightarrow \lambda = 1 + \frac{R}{100}$$

and

$$A(n)_{\text{CF}} = A\left(1 + \frac{R}{100}\right)^n.$$

$$A(n)_{\text{PI}} = C, \text{ where } C = \frac{-P}{1 - (1 + \frac{R}{100})}$$

(The power terms cancel out each other due to the coefficient of $A(n)$. Therefore we only take the coefficient of $A(n)$ and $A(n+1)$.) And so

$$A(n) = A\left(1 + \frac{R}{100}\right)^n - \frac{P}{\frac{-R}{100}}$$

We also need to choose appropriate A so that initial balance is $A(0)$.

Note: The methods employed in the previous exmaples are just like those we used for differential equations which have the property of linearity.

General Case with constant coefficients

$$\begin{aligned}\mathscr{L}U(n) = a_0U(n+m) + a_1U(n+m-1) + a_2U(n+m-2) + \cdots \\ + a_{m-1}U(n+1) + a_mU(n) = f(n)\end{aligned}$$

with $a_0, a_1, \ldots, a_m$ constants. The equation is linear, order m. It is homogeneous iff $f(n) = 0$, and inhomogeneous iff $f(n) \neq 0$.

The General Solution (GS) can always be written as

$$U_{\text{GS}} = U_{\text{CF}} + U_{\text{PI}}$$

where $\mathscr{L}U_{\text{CF}} = 0$, $\mathscr{L}U_{\text{PI}} = f(n)$. U_{CF} has m arbitrary constants, while U_{PI} is any solution i.e. it is not unique.

For the CF with a constant coefficient equation we try $U(n)_{\text{CF}} \propto \lambda^n$

$$\Rightarrow \lambda^n[a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_{m-1}\lambda + a_m] = 0$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are roots of this characteristic equation. Then

$$U(n)_{\text{CF}} = A_1\lambda_1^n + A_2\lambda_2^n + \cdots + A_m\lambda_m^n$$

with $A_1, A_2, \ldots, A_m$ being arbitrary constants.

Example 65.

(1)
$$U(n+2) + 7U(n+1) - 18U(n) = 0$$
$$\Rightarrow \lambda^2 + 7\lambda - 18 = 0, \lambda_1 = -9, \lambda_2 = 2.$$
$$\Rightarrow U(n) = A_1(-9)^n + A_2(2)^n.$$

What about the equal roots case?

(2)
$$U(n+2) - 6U(n+1) + 9U(n) = 0$$
$$\Rightarrow \lambda^2 - 6\lambda + 9 = 0, \lambda_1 = \lambda_2 = 3.$$

Certainly we have $A_1(3)^n$, but we need something else! — It is $A_2 n(3)^n$.

$$\Rightarrow U(n) = A_1 3^n + A_2 n 3^n.$$

What about a PI? Well, as for differential equations, it all depends on $f(n)$!

(a) $f(n) = Cp^n$ where $p \neq \lambda_1$ or λ_2, and C is a constant.

This is easy! $U(n)_{\text{PI}} = Ap^n$ with A chosen suitably. From our earlier example, we put

$$U(n+2) + 7U(n+1) - 18U(n) = 6(4)^n.$$

Since $4 \neq -9$ or 2 we can write $U(n)_{\text{PI}} = A(4^n)$,

$$A(4^{n+2}) + 7A(4^{n+1}) - 18A(4^n) = 6(4^n)$$

i.e. $16A + 28A - 18A = 6 \Rightarrow A = \frac{3}{13}$. So

$$U_{\text{GS}} = A_1(-9)^n + A_2(2)^n + \frac{3}{13}(4)^n.$$

(b) $f(n) = Cp^n$ where $p = \lambda_1$ (say)

Just as for a differential equations we need a more complicated $U(n)_{\mathrm{PI}} = A(n)\lambda_1^n$, where $A(n)$ is a polynomial in n. Again from out earlier example, we put

$$U(n+2) + 7U(n+1) - 18U(n) = 3(2)^n.$$

Let's say

$$U(n)_{\mathrm{PI}} = A(n)(2)^n = (a + bn + cn^2)(2^n)$$

Well, apparently $a = 0$, after comparing with $U(n)_{\mathrm{CF}}$. Then

$$\begin{aligned}[b(n+2) + c(n+2)^2]2^{n+2} &+ 7[b(n+1) + c(n+1)^2]2^{n+1} \\ &- 18(bn + cn^2)2^n = 3(2^n)\end{aligned}$$

Cancel a factor of 2^n, then the n^2 terms are cancelled, and n terms leave $4(b+4c) + 14(b+2c) - 18b = 0$, and constant terms leave $4(2b + 4c) + 14(b+c) = 3$.

$$\Rightarrow c = 0 \text{ and } b = \frac{3}{22}.$$

So

$$U_{\mathrm{GS}} = A_1(-9)^n + A_2(2)^n + \frac{3}{22}n(2^n)$$

and so on...

Since our equation is linear, we can just add terms together to construct $U(n)_{\mathrm{PI}}$ for quite complicated $f(n)$ on RHS.

Some results can seem very strange! The Binet formula for Fibonacci numbers involved irrational numbers as building blocks — but produced integers!

<u>Example</u>:

$$U(n+2) - 2U(n+1) + 5U(n) = 0$$

with say $U(1) = 6, U(2) = 2$ (so that $U(0) = 2$) which obviously produces a sequence of integers. However,

$$\lambda^2 - 2\lambda + 5 = 0 \Rightarrow \lambda_1 = 1 + 2i, \lambda_2 = 1 - 2i.$$

So

$$U(n) = A_1(1+2i)^n + A_2(1-2i)^n$$

Substitute $n = 0, 1$ into the equation, and we get

$$A_1 = 1 - i, A_2 = 1 + i$$

and

$$U(n) = (1-i)(1+2i)^n + (1+i)(1-2i)^n.$$

So $U(3) = -26$, etc.

(c) $f(n)$ is a polynomial in n

Well here we just need to choose a suitable polynomial and choose the coefficients to fit the case.

Example: Try to find

$$S(n) = 1^2 + 2^2 + \cdots + n^2 = \sum_{r=1}^{n} r^2.$$

If we knew the answer or could guess, then we could confirm using induction. If not we can just recognize that

$$S(n+1) - S(n) = (n+1)^2$$

We can easily see that $\lambda = 1$, implying that

$$S(n)_{\mathrm{CF}} = A(1)^n = A.$$

Then

$$S(n)_{\mathrm{PI}} = an^3 + bn^2 + cn.$$

(Do not need a constant term here since it is already in CF.) So

$$a(n+1)^3 + b(n+1)^2 + c(n+1) - an^3 - bn^2 - cn = (n+1)^2$$

Comparing the coefficients, we get $a = \frac{1}{3}, b = \frac{1}{2}, c = \frac{1}{6}$. So

$$S(n)_{\mathrm{GS}} = A + \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

and $A = 0$ since we know $S(0) = 0, S(1) = 1$, etc. So

$$S(n) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n = \frac{1}{6}n(n+1)(2n+1).$$

This method is constructive, and we can extend the idea to find $\sum_{r=1}^{n} r^3 = \left[\frac{1}{2}n(n+1)\right]^2$, etc.

As always, if we tried a polynomial PI which is too simple, or too complicated, the calculation is self-correcting!

(d) $f(n) = (\text{polynomial in } n)(p)^n$

Just like our previous cases our expectation is

$$U(n)_{\text{PI}} = (\text{suitable polynomial})(p)^n.$$

Then following similar step: matching coefficients, substitute in values, obtain value of the constant if boundary condition is provided, etc.

2.2.3 Differencing and Difference Tables

Definition 66. The (forward) ***difference operator*** Δ is defined by

$$\Delta U(n) = U(n+1) - U(n)$$

so that

$$\begin{aligned}\Delta^2 U(n) &= \Delta[U(n+1) - U(n)] \\ &= \Delta U(n+1) - \Delta U(n) \\ &= [U(n+2) - U(n+1)] - [U(n+1) - U(n)] \\ &= U(n+2) - 2U(n+1) + U(n)\end{aligned}$$

Now we can see that $\Delta n^k = (n+1)^k - n^k = kn^{k-1} + \cdots + 1$, and this means that

$$\Delta(\text{polynomial in } n \text{ of degree } k) = (\text{polynomial in } n \text{ of degree } (k-1))$$

We can continue this process of course, $\Delta(\Delta(\Delta(\ldots))) = \Delta^k()$.

$$\Rightarrow \Delta^k(\text{polynomial of degree } k) = (\text{polynomial of degree } 0)$$

and $\Delta^{k+1}(\text{polynomial of degree } k) = 0$.

Note: Successive differencing is a *discrete* analogy to differentiation. ($\frac{\mathrm{d}^4}{\mathrm{d}x^4}(x^4) = 24$ of course!) We can consider the reverse (*inverse*) of the differencing process ($\approx$ integration).

Example 67.

$$
\begin{aligned}
\Delta n^4 &= (n+1)^4 - n^4 = 4n^3 + 6n^2 + 4n + 1 \\
\Delta^2(n^4) &= \Delta(4n^3) + \Delta(6n^2) + \Delta(4n) + \Delta(1) = 12n^2 + 24n + 14 \\
\Delta^3(n^4) &= 24n + 36 \\
\Delta^4(n^4) &= 24 \\
\Delta^5(n^4) &= 0.
\end{aligned}
$$

Chapter 3

Linear Algebra

3.1 Introduction to Matrices and Vectors

3.1.1 Column vectors

Definition 68. A *column vector* (n-column vector) $\boldsymbol{v}_n$ is a tuple of n real numbers written as a single columnn, with $a_1, a_2, a_3, \ldots, a_n \in \mathbb{R}$:

$$\boldsymbol{v}_n := \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}$$

Definition 69. $\mathbb{R}^n$ is the set of all column vectors of height n whose entries are real numbers. In symbols:

$$\mathbb{R}^n = \{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} : a_1, a_2, \ldots, a_n \in \mathbb{R}\}$$

Example 70. $\mathbb{R}^2$ can be seen as Euclidean plane. $\mathbb{R}^3$ can be seen as Euclidean space.
Caution: Our vectors always “start” at the origin.

Definition 71. The ***zero vector*** $\mathbf{0}_n$ is the height n-column vector all of whose entries are 0.

Definition 72. The ***standard basis vectors*** in $\mathbb{R}^n$ are the vectors

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

i.e. $\boldsymbol{e}_k$ is the vector with kth entry equal to 1 and all other entries equal to 0.

Operations on column vectors

$$\boldsymbol{v} := \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \quad \boldsymbol{u} := \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

be column vectors $\mathbb{R}^n$, and let λ be a (real or complex) number.

(1) Addition on vectors in $\mathbb{R}^n$ is given by:

$$\begin{pmatrix} v_1 + u_1 \\ v_2 + u_2 \\ \vdots \\ v_n + u_n \end{pmatrix}$$

$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ (binary operatioin). $(\mathbb{R}^n, +)$ is a group.

(2) ***Scalar multiplication*** $\lambda \boldsymbol{v}$ on $\mathbb{R}^n$:

$$\begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}$$

$s : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, so not binary operation.

(3) ***Dot product*** $v \cdot u$ is defined to be the number $v_1u_1 + v_2u_2 + \cdots + v_nu_n$. $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, so not binary.

Example 73. Show that $(\mathbb{R}^n, +)$ is an Abelian group.

- Identity: $\mathbf{0}_n$ $(v + \mathbf{0}_n = \boldsymbol{v})$
- $-\boldsymbol{v}$ are inverses, where
$$-\boldsymbol{v} := \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix}$$
- associativity: $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$.
- commutative: $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

Caution: $+$ only makes sense for vectors of the *same size*. e.g. $\boldsymbol{v} \cdot \mathbf{0}_n = 0 \in \mathbb{R}$.

Definition 74. let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n, \lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n \in \mathbb{R}$, then
$$\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_n\boldsymbol{v}_n$$
is called a ***linear combination*** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \ldots, \boldsymbol{v}_n$.

Definition 75. The set of all linear combinatioins of a collection of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ is called the ***span*** of the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$.
Notation:
$$\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} := \{\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_n\boldsymbol{v}_n | \lambda_1, \ldots, \lambda_n \in \mathbb{R}\}$$

Example 76. compute the span of

- $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$, $\boldsymbol{e}_1, \boldsymbol{e}_2 \in \mathbb{R}^2$.
$$\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \{\lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 | \lambda_1, \lambda_2 \in \mathbb{R}\} = \{\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} | \lambda_1, \lambda_2 \in \mathbb{R}\}$$
- $\operatorname{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}\right\} = \{\begin{pmatrix} \lambda_1 \\ 2\lambda_2 \\ 0 \end{pmatrix} | \lambda_1, \lambda_2 \in \mathbb{R}\}$

Definition 77. let $\boldsymbol{v} \in \mathbb{R}^n$. The ***length*** of $\boldsymbol{v}$, a.k.a. the ***norm*** of $\boldsymbol{v}$, is the non-negative real number $\|\boldsymbol{v}\|$ defined by

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

Note: $\|\mathbf{0}\| = 0$, and conversely if $\boldsymbol{v} \neq 0$ then $\|\boldsymbol{v}\| > 0$. This definition agrees with out usual ideas about the length of a vector in $\mathbb{R}^2$ or $\mathbb{R}^3$, which follows from Pythagoras' theorem.

Definition 78. A vector $\boldsymbol{v} \in \mathbb{R}^n$ is called a ***unit vector*** if $\|\boldsymbol{v}\| = 1$.

Example 79.

(1) Any non-zero vector $\boldsymbol{v}$ can be made into a unit vector $\hat{\boldsymbol{u}} := \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$. This process is called ***normalizing***.

(2) The standard basis vectors are unit vectors.

3.1.2 Basic Matrix Operations

Definition 80. An $n \times m$-matrix is a rectangular grid of numbers called the *entries* of the matrix with n rows and m columns. A real matrix is onne whose entries are real numbers, and a complex matrix is one whose entries are complex numbers.

Notations: $M_{n\times m}(\mathbb{R}), M_{n,m}(\mathbb{R}), \mathrm{Mat}_{n\times m}(\mathbb{R}), \mathbb{R}^{n\times m}$.

Operations on matrices:

Definition 81. let $A = (a_{ij})$ and $B = (b_{ij})$ are $n \times m$-matrix, $\lambda \in \mathbb{R}$. Then:

(1) $A + B = n \times m$-matrix $(a_{ij} + b_{ij})$. $+ : M_{n\times m}(\mathbb{R}) \times M_{n\times m}(\mathbb{R}) \to M_{n\times m}(\mathbb{R})$

(2) $\lambda A = n \times m$-matrix (λa_{ij})

Theorem 82. $(M_{n\times m}(\mathbb{R}), +)$ is an Abelian grouop.

Definition 83. The ***transpose*** A^T of an $n\times m$-matrix (a_{ij}) is the $m\times n$-matrix (a_{ij}). The ***leading diagonal*** of a matrix is the $(1,1),(2,2),\dots$ entries. So the transpose is obtained by doing a reflection in the leading diagonal.

(Multiplying matrices with vectors) Definition 84. Let $A = (a_{ij})$ be an $n \times m$-matrix, $\boldsymbol{v} \in \mathbb{R}^m$. Then $A\boldsymbol{v}$ is the vector in $\mathbb{R}^n$ with i-th row entry $\sum_{j=1}^m a_{ij}\boldsymbol{v}_j$

Example 85.

- Prove that for $A \in M_{n\times m}(\mathbb{R}), \boldsymbol{e}_k \in \mathbb{R}^m$, $A\boldsymbol{e}_k = k$-th column of A.

  Proof: let $A = (a_{ij})$. By definition the i-th entry of $A\boldsymbol{e}_k$ is

  $$\sum_{j=1}^{m} a_{ij}(\boldsymbol{e}_k)_j = a_{ik}$$

  since $(\boldsymbol{e}_k)_j = 0$ whenever $j \neq k$, 1 for $j = k$

- Let I_n be the identity matrix. Show formally that $I_n\boldsymbol{\nu} = \boldsymbol{\nu}$, $\forall \boldsymbol{\nu} \in \mathbb{R}^n$.

- $\boldsymbol{\nu} \cdot \boldsymbol{v} = \boldsymbol{\nu}^T\boldsymbol{v}$

- let $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \boldsymbol{\nu}_3 \in \mathbb{R}^3$. Write the linear combination $3\boldsymbol{\nu}_1 - 5\boldsymbol{\nu}_2 + 7\boldsymbol{\nu}_3$ as a multiplication of matrix $A \in M_{3\times 3}(\mathbb{R})$ with a vector $\boldsymbol{x} \in \mathbb{R}^3$. Then

  $$A\boldsymbol{x} = \begin{pmatrix} \boldsymbol{\nu}_1 & \boldsymbol{\nu}_2 & \boldsymbol{\nu}_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1\boldsymbol{\nu}_1 + x_2\boldsymbol{\nu}_2 + x_3\boldsymbol{\nu}_3$$

  with $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \boldsymbol{\nu}_3$ written as a column vector to form a matrix in the above expression, thus using matrix multiplication to express linear combination of vectors.

3.2 Systems of linear equations

Definition 86. A ***linear equation*** in the variables $x_1, x_2, \ldots, x_n \in \mathbb{R}$ is an equation of the form:

$$\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n = c, \text{ with } \lambda_1, \ldots, \lambda_n \subset \textit{ Fixed} \text{ real numbers}$$

Caution: In particular, no powers/multiplications/function of one or more variables.

Definition 87. A system of n linear equations is a list of simultaneous linear equations. It can be converted to $A\boldsymbol{x} = \boldsymbol{b} \in \mathbb{R}^m$,with

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

Caution: Thee $m \times n$-matrix A is called coefficient matrix. The matrix $(A|\boldsymbol{b})$ where the vector $\boldsymbol{b}$ is added as a column on the right is called ***augmented matrix***.

Definition 88. A system is called ***consistent*** (resp. inconsistent) if it has a solution $(s_1, s_2, \ldots, s_m)$ (resp. no solution).

Example 89.

$$\begin{cases} x_1 + x_3 - x_4 = 1 \\ x_2 - x_4 = 6 \\ x_1 + x_2 + 6x_3 - 3x_4 = 0 \end{cases}$$

Augmented matrix form:

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & -1 & 6 \\ 1 & 1 & 6 & -3 & 0 \end{array}\right)$$

Definition 90. A ***row operation*** is one of the following procedures on a $n \times m$-matrix (a_{ij}):

(1) $r_i(\lambda)$: multiply row i by a scalar $\lambda \in \mathbb{R}, \lambda \neq 0$.

(2) r_{ij}: swap row i with row j.

(3) $r_{ij}(\lambda)$: multiply row i by $\lambda \neq 0$, $\lambda \in \mathbb{R}$ and add it to row j.

Example 91. let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, so

$$r_{12} \Rightarrow \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$$

$$r_2(2) \Rightarrow \begin{pmatrix} 1 & 2 \\ 6 & 8 \end{pmatrix}$$

$$r_{12}(2) \Rightarrow \begin{pmatrix} 1 & 2 \\ 5 & 8 \end{pmatrix}$$

Proposition 92. Let $A\boldsymbol{x} = \boldsymbol{b}$ be a system of linear equations in matrix form, $(A|\boldsymbol{b})$ the augmented matrix, $(A'|\boldsymbol{b}')$ the augmented matrix of the system after row operation. Show that x is solution of $A\boldsymbol{x} = \boldsymbol{b} \iff x$ is solution of $A'\boldsymbol{x} = \boldsymbol{b}'$.

Proof. row operations of type (1) and (2) $\Rightarrow$ trivial.

(3) Take equation i, multiply it by λ, add it to equation j. $\Rightarrow (a_{j1} + \lambda a_{i1})x_1 + \cdots + (a_{jm} + \lambda a_{im})x_m = b_j + \lambda b_i$. □

Caution: Every row operation is invertible:

$$[r_i(\lambda)]^{-1} = r_i(\frac{1}{\lambda}),\ [r_{ij}]^{-1} = r_{ij},\ [r_{ij}(\lambda)]^{-1} = r_{ij}(-\lambda)$$

3.2.1 Gauss algorithm

Definition 93. The left most non-zero entry in a non-zero row is called ***leading entry***. A matrix is called in ***echelon form*** if:

(1) The leading entry in each non-zero row is 1.

(2) The leading 1 of each row is to *the right* of the leading 1 in the row above.

(3) The zero-rows are *below* all other rows.

Example 94.

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Only the last one is in echelon form.

Definition 95. A matrix is ***row reduced echelon form*** if:

(1) It is in echelon form.

(2) The leading 1 in each row is the *only* non-zero entry in its column.

Example 96.

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & \alpha & \beta & 2 \\ 0 & 0 & 1 & -2 \end{pmatrix}.$$

The second one is not.

Chapter 4

Analysis